

BEST AVAILABLE COPY

## Protein domain screening system for structural genomics

○Eiko Seki<sup>1</sup>, Takanori Kigawa<sup>1,2</sup>, Natsuko Matsuda<sup>1</sup>, Yoshihide Hayashizaki<sup>3</sup>, Shigeyuki Yokoyama<sup>1,2,4</sup>

<sup>1</sup>*Protein Research G., GSC, RIKEN Yokohama Institute*, <sup>2</sup>*Cell Signal Lab., RIKEN Harima Institute at Spring-8*, <sup>3</sup>*Genome Exploration Research G., GSC, RIKEN Yokohama Institute*, <sup>4</sup>*Grad. Sch. of Sci., Univ. of Tokyo*

In RIKEN GSC Protein Research Group, we are heading for the systematical determination of the three-dimensional structures of protein folds. For this purpose, it is essential to prepare soluble and properly folded protein samples for structural analyses.

Here, we constructed a new experimental system, which enables us to select suitable samples for structural analyses. To begin with, target cDNA was randomly fragmented with Exonuclease III, to construct a deletion library of various lengths. For the 1st step screening, deleted fragments were expressed as C-terminal GFP-fusions in *E.coli*, and fluorescing clones were selected as candidates. For the 2nd step screening, those selected clones were expressed in cell free synthesis system, and screened on the basis of the intensity of GFP fluorescence.

We verified this screening system with mouse growth factor receptor-binding protein 2 (Grb2). The GFP fluorescence in the 2nd step screening was correlated with the solubility of deleted fragments. The terminals of soluble fragments coincided with the structural boundaries of Grb2. Thus, by using this system, we could screen protein samples for structural analyses, from the pool of various proteins. We are now applying this system to define structural domains of mouse cDNA.

## 序論

構造ゲノム科学は、タンパク質の立体構造解析を系統的かつ網羅的に行うものである。数万個といわれるタンパク質の立体構造は、数千種類の基本構造の組み合わせから形成されていると考えられており、我々はタンパク質の基本構造を網羅的に解明することにより、タンパク質の立体構造の全貌を明らかにしようとしている。この目的のためには、NMRやX線結晶構造解析を用いた立体構造解析に適した、構造をとり、かつ可溶性の高いタンパク質試料を、迅速に選び出す系の構築が重要である。

Waldo (1999) らの報告によると、タンパク質のC末端側にGFPを融合させた状態で大腸菌内で発現させることにより、その可溶性を検出することが可能である。しかしながら、大腸菌内で発現したタンパク質の蛍光強度を測定することは、煩雑であり、多検体同時処理に向いているとはいえない。そこで、我々は、cDNA deletion libraryのGFP-fusionと無細胞タンパク質合成系を組み合わせることにより、より正確かつ簡便に、可溶性のタンパク質ドメインを選び出す系の構築を行った。

## タンパク質ドメイン選択法の開発

- 実験的タンパク質ドメイン選択法の必要性
  - 立体構造解析において、構造をとり、かつ可溶なタンパク質試料が必要である。
  - タンパク質の可溶性は、一次配列上からは予測出来ず、実験的手法でしか確認できない。
  - 構造ゲノム科学では多検体同時処理による構造決定が必要である。
- 必須条件
  - 簡便かつ迅速な方法
  - 立体構造をとったドメインの選択
  - 可溶性ドメインの選択
  - 大量合成系への簡便な移行

## GFP融合体 を利用した 可溶性サンプルのスクリーニング

### GFP 融合体:

可溶性タンパク質をC末端側でGFP融合体にすると発光する

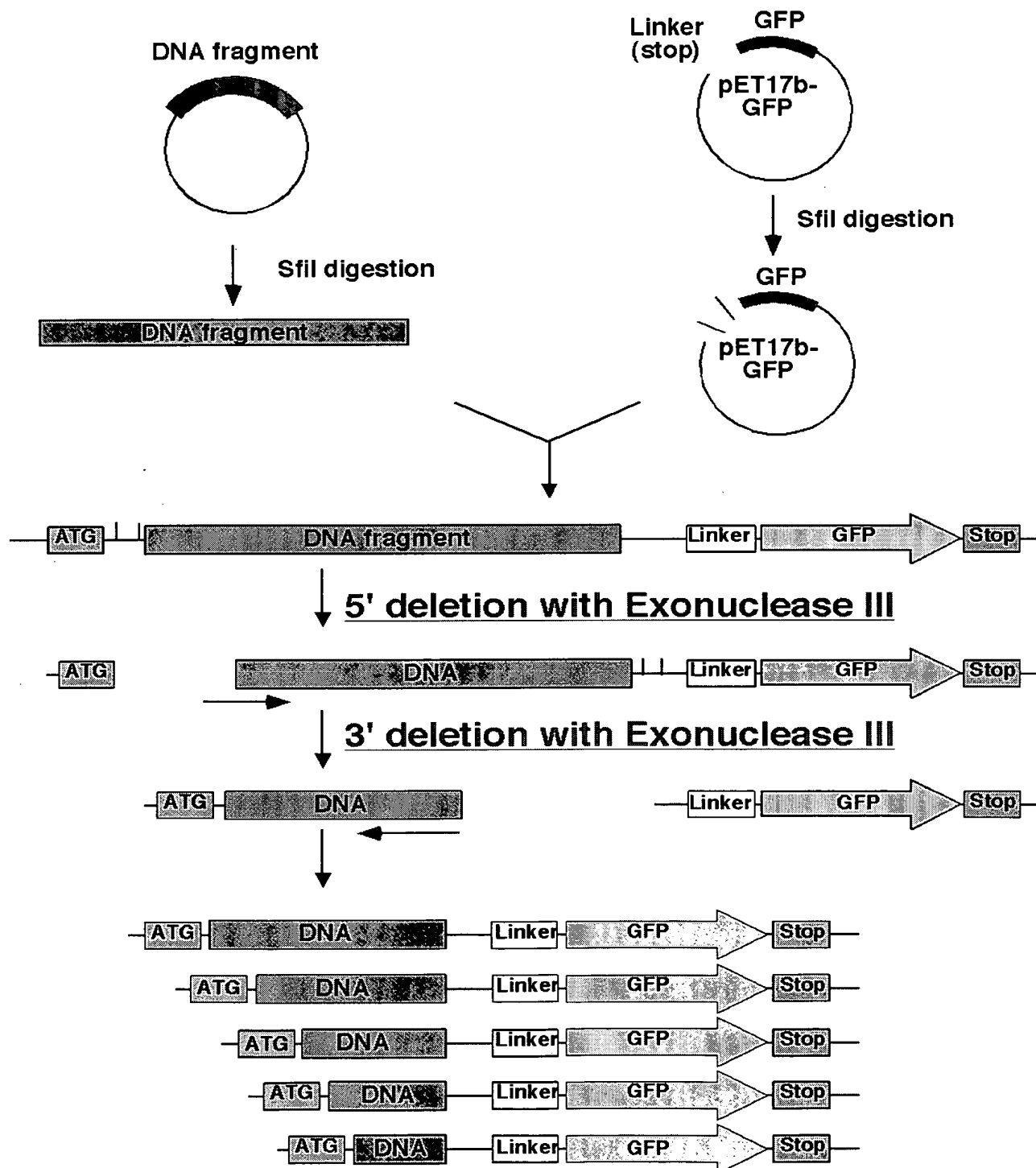
⇔ 難溶性タンパク質は発光しない

(Waldqvist G S et al Nat Biotechnol 17, 691-695 (1999))



様々な長さを含む均質で良質なdeletion libraryと組み合わせることで、効率よく構造解析に適したタンパク質ドメインをスクリーニングできる

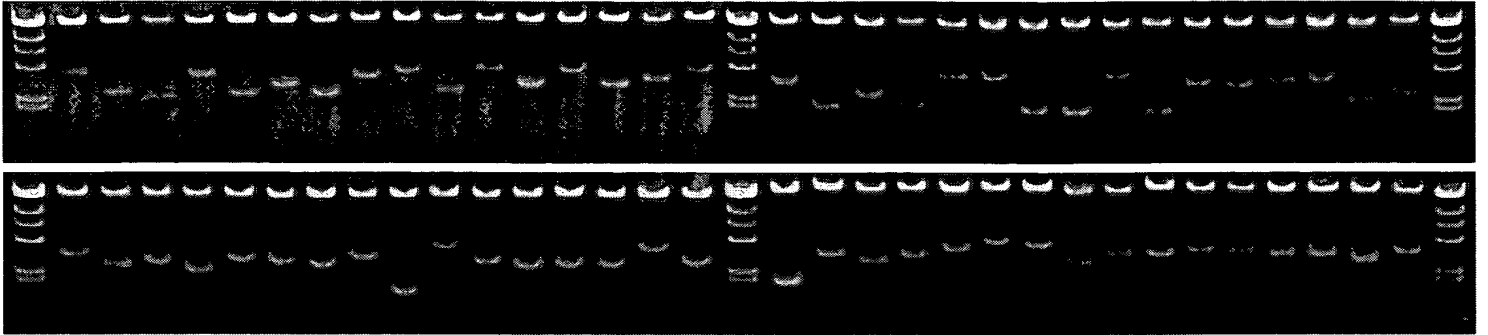
# Exonuclease IIIを用いた Deletion Libraryの作成方法



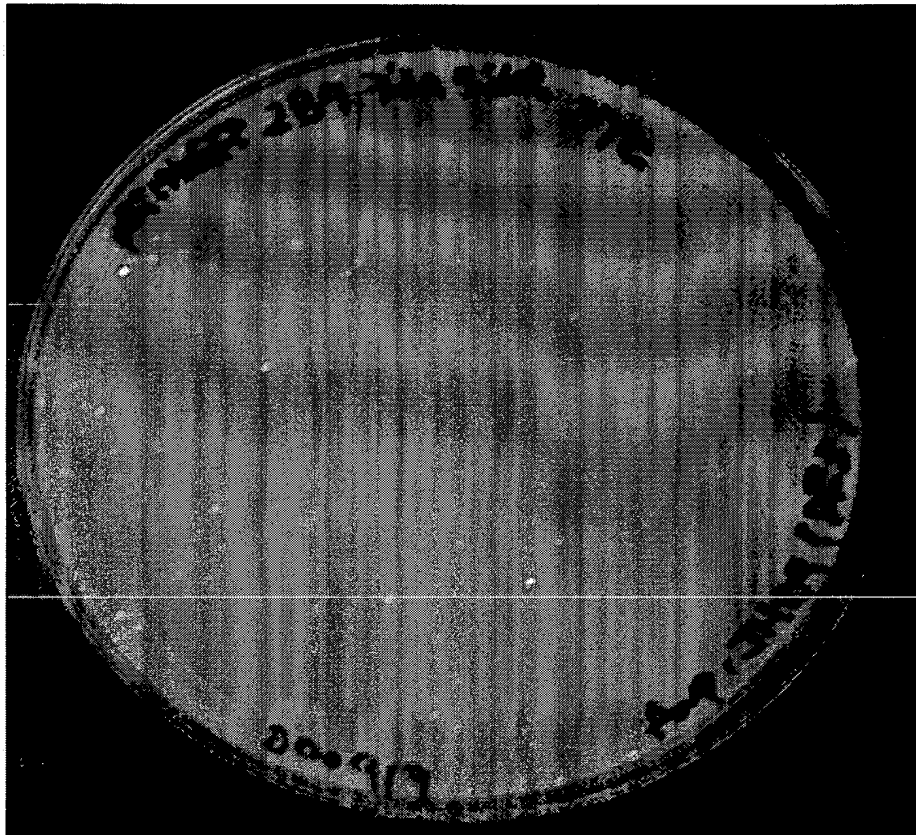
# Target DNAの制限酵素サイト検索

Trial	OK	2enz change	No enz.	O: Zeislos deletion	Δ: 1 side deletion	X: no deletion																			
	General ID	P.P. rio	MW [Da]	PCR成功度	R.T.2	R.SUP	Pat1 cut	EcoR1 cut	Sma1 cut	Kpn1 cut	Spe1 cut	Nhe1 cut	J' Sph1	3'Nsi1	5'Cla1	5'Kai1	EcoR1 V	Nru1	Pml1	Stu1	SacII	NarI	HindIII	NotI	Xba1
X	14 1B2	76	ZX00001G13A	10740	3	488.9	411																		
Δ	161 1F75	70	ZX00003K23A	16300	3	157.3	94.87	214	###	###	###	###	###	292	###	###	###	###	###	###	###	###	41	###	###
X	2 1A2	66	ZX00001E07A	17121	3	360.9	245.8	238	1	###	###	###	###	###	###	###	###	###	###	###	###	###	19	###	###
X	96 1H12	36	ZX00002H05A	17206	3	187.9	152.3	1331	587	1071	###	208	###	1638	###	###	###	###	###	25	###	###	###	###	###
Δ	116 2B8	1	ZX00003M12A	20830	3	325.1	119.3	310	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	66
X	70 1F10	38	ZX00002K10A	22088	3	534.1	96.9	258	###	###	###	###	###	600	###	###	###	###	###	###	###	1309	###	894	
X	38 1D2	59	ZX00001I07A	25649	3	424	303.1	1404	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
○/Δ	144 1012	64	ZX00003L10A	29655	3	440.2	129.5	53	###	###	###	61	###	###	###	###	496	###	###	###	###	41	###	###	###
○	56 1E8	48	ZX00002E08A	32885	3	215.6	202.8	###	7	###	###	###	###	###	###	###	###	###	###	282	###	015	###	###	###
○/Δ	48 1D10	40	ZX00002G14A	33400	3	366.7	131.6	328	###	###	###	###	###	682	###	###	###	###	###	###	###	###	###	###	###
○	187 2H7	23	ZX00003M04A	35674	3	288	216.5	706	###	452	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
○/Δ	191 2H11	62	ZX00003H22A	38197	3	25.68	18.69	###	###	###	398	###	###	###	###	###	###	###	###	302	###	1285	###	###	###
○	122 1C3	29	ZX00002I01A	39769	3	193.1	153.3	815	350	###	###	###	###	###	###	###	###	###	###	###	###	57	###	###	###
000201	○	184 2H4	72	ZX00003G09A	42083	3	57.88	50.12	25	1472	###	###	###	288	908	1725	###	###	###	###	###	1177	###	###	###
000201	○	192 2H12	71	ZX00003F02A	43663	3	236.9	51.23	222	###	###	783	783	###	20	###	703	403	###	###	###	703	1303	###	537
○	173 2G5	54	ZX00003M07A	44603	3	14.93	13.4	802	###	472	687	###	###	###	###	###	###	###	567	###	717	###	1014	###	###
○/Δ	118 2B11	42	ZX00003N21A	47702	3	134.5	87.5	###	###	###	138	###	###	999	###	###	###	###	###	1075	###	1450	###	812	###
○	73 1G1	21	ZX00001C13A	48119	3	390.1	320.1	635	###	###	659	###	###	###	###	###	###	###	###	###	###	###	###	###	86
○	62 1F2	7	ZX00001K09A	50436	3	303.9	197.9	725	###	###	1221	###	###	533	983	1255	61	###	###	327	###	61	524	29	###
○	181 2H1	27	ZX00002P05A	52085	3	26.74	23.77	76	###	79	76	###	336	937	###	###	###	###	###	###	###	679	###	###	###
○/Δ	143 2D1	77	ZX00003F02A	52580	3	152.4	116.1	1444	###	10	###	###	###	5245	###	###	288	###	###	###	###	1466	###	###	###
○/Δ	75 1G3	32	ZX00001C08A	54264	3	440.2	384.4	140	###	612	89	###	951	1249	###	###	###	###	###	932	###	1970	###	###	###
○/Δ	7 1A7	47	ZX00001F14A	60593	3	482.1	77.87	1391	70	1193	###	941	###	###	###	###	###	###	###	###	###	1076	###	###	###
○	166 2F10	24	ZX00002L13A	61309	3	111.3	96.99	50	525	###	290	270	###	87	###	1421	###	###	###	976	###	1456	###	###	###
○	111 2B5	74	ZX00002J24A	64780	3	20	14.35	898	###	1085	###	36	543	1570	1140	###	###	###	705	1322	###	705	###	###	###
000201	○	133 2D1	9	ZX00002J21A	65572	3	67.48	46.15	2165	2676	###	1829	###	###	###	1735	583	###	2232	###	###	1927	###	###	###
○	76 1B4	4	ZX00001B03A	69242	3	288.5	190.7	1653	1156	###	###	120	120	###	1778	###	###	###	###	27	###	1142	###	###	###
○	112 2B4	14	ZX00003A21A	82257	3	191.1	123.2	6	###	###	###	1804	###	2502	###	###	###	###	865	###	###	###	###	###	1635
○/Δ	27 1G3	73	ZX00001M06A	20905	2	43.03	26.17	158	###	416	###	###	###	###	###	###	###	###	###	###	###	###	###	###	1002
X	67 1C7	44	ZX00001L04A	30333	2	268.3	171.9	838	###	###	###	324	###	###	###	###	###	###	###	###	1668	###	###	2406	
X	89 1F8	28	ZX00002L01A	31739	2	88.50	50.78	31	###	###	###	###	###	###	535	554	###	###	###	###	1279	###	###	###	
X	55 1E5	53	ZX00001J05A	49239	2	132.5	87.51	397	###	###	###	###	###	###	997	###	###	###	###	###	###	###	###	###	###
Δ	82 1H8	60	ZX00002E23A	49556	2	188.5	163.4	588	###	447	1326	###	###	1424	###	###	###	###	333	###	###	###	###	###	###
○	21 1H9	20	ZX00003J11A	52930	2	85.19	139.9	229	###	1064	23	159	###	###	881	###	###	###	692	###	18	###	###	###	###
○/Δ	85 1H4	57	ZX00001B21A	71175	2	81.99	60.08	155	###	1052	78	###	###	1870	###	###	###	###	96	###	###	1258	###	###	1455
○	1 1A1	56	ZX00001A05A	76132	2	107.3	62.98	56	###	1205	230	###	1698	###	1818	1395	###	733	813	1053	1395	###	###	###	
X	129 2C8	79	ZX00003F19A	13178	1	304.5	124.1	360	###	###	###	15	###	###	110	###	###	###	###	###	###	###	###	###	###
X	183 2H3	55	ZX00002P10A	15051	1	280.9	68.85	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	546
X	176 2G8	45	ZX00003F11A	24914	1	335.5	118.3	722	58	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
○/Δ	137 2D5	39	ZX00003J15A	40009	1	216.5	136.3	509	###	928	###	###	###	###	661	###	###	###	###	###	###	1138	###	###	###
○	177 2G9	49	ZX00003J09A	45841	1	188.5	175.7	###	###	1056	338	###	###	515	###	687	###	###	128	###	687	###	###	1648	
○	124 2C4	52	ZX00003C03A	46051	1	216.9	137.6	621	###	###	###	###	###	398	###	1117	886	###	###	###	804	###	###	###	
○	61 1F1	46	ZX00001C01A	51615	1	554.6	260.4	408	###	###	###	###	###	###	###	529	###	###	374	640	###	###	###	###	###
○	140 2D8	25	ZX00003O20A	54874	1	390.4	188.1	###	817	###	###	###	###	###	###	529	###	###	###	###	529	366	###	###	###
○	190 2H10	13	ZX00003N09A	61565	1	303.7	115.5	875	207	###	1535	###	###	1098	###	49	###	1201	###	1299	###	###	###	1455	###

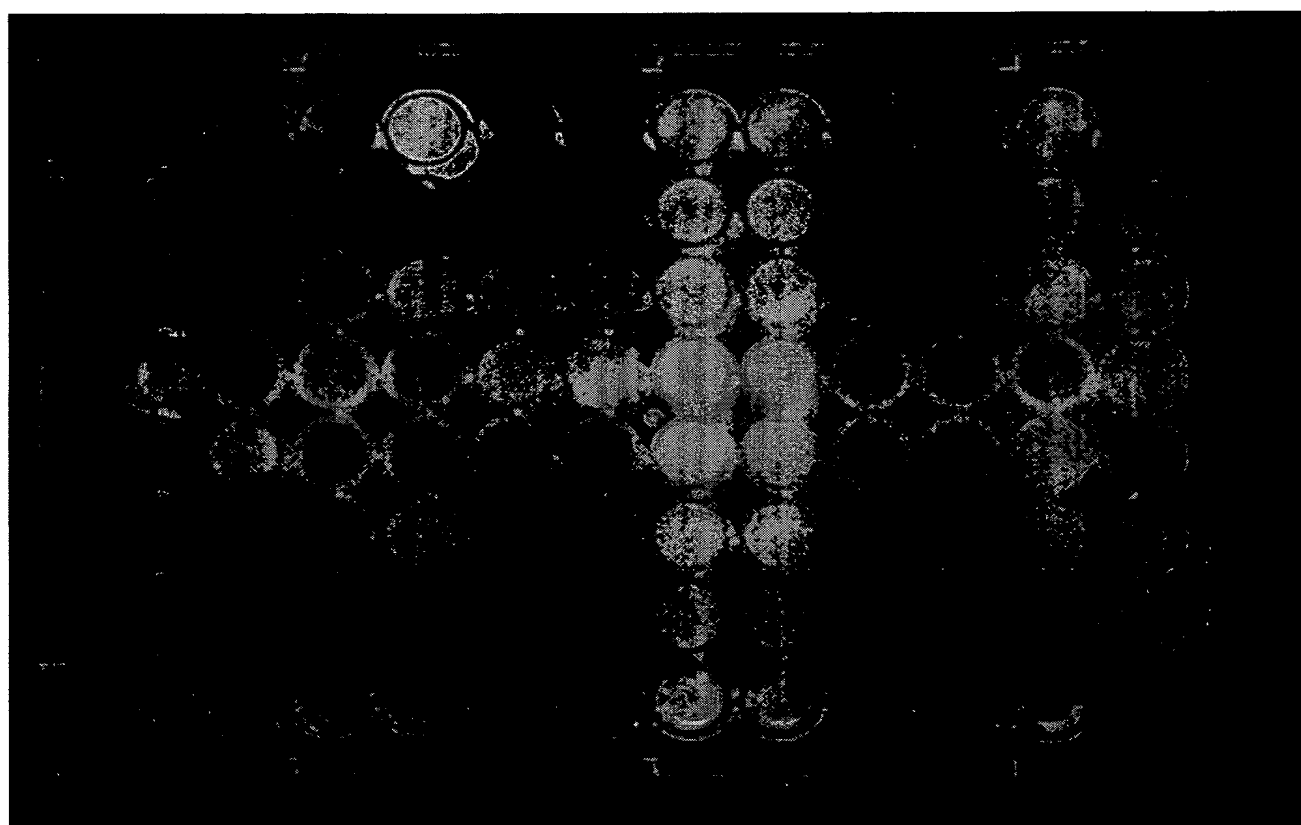
## 作成したDeletion Library



## 大腸菌での1段階目スクリーニング

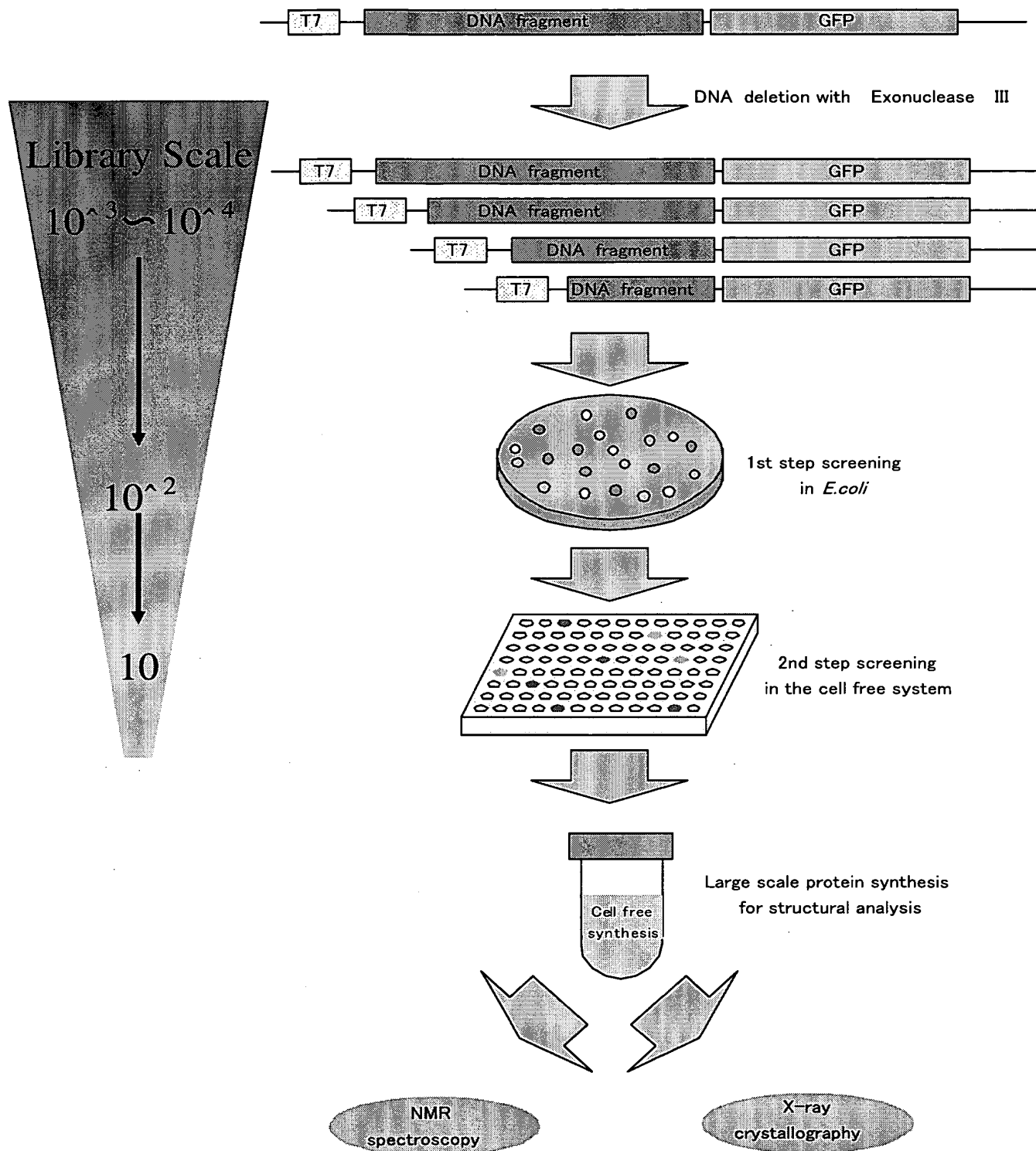


# 無細胞合成での 2段階目スクリーニング





# タンパク質ドメインの スクリーニングシステムの流れ

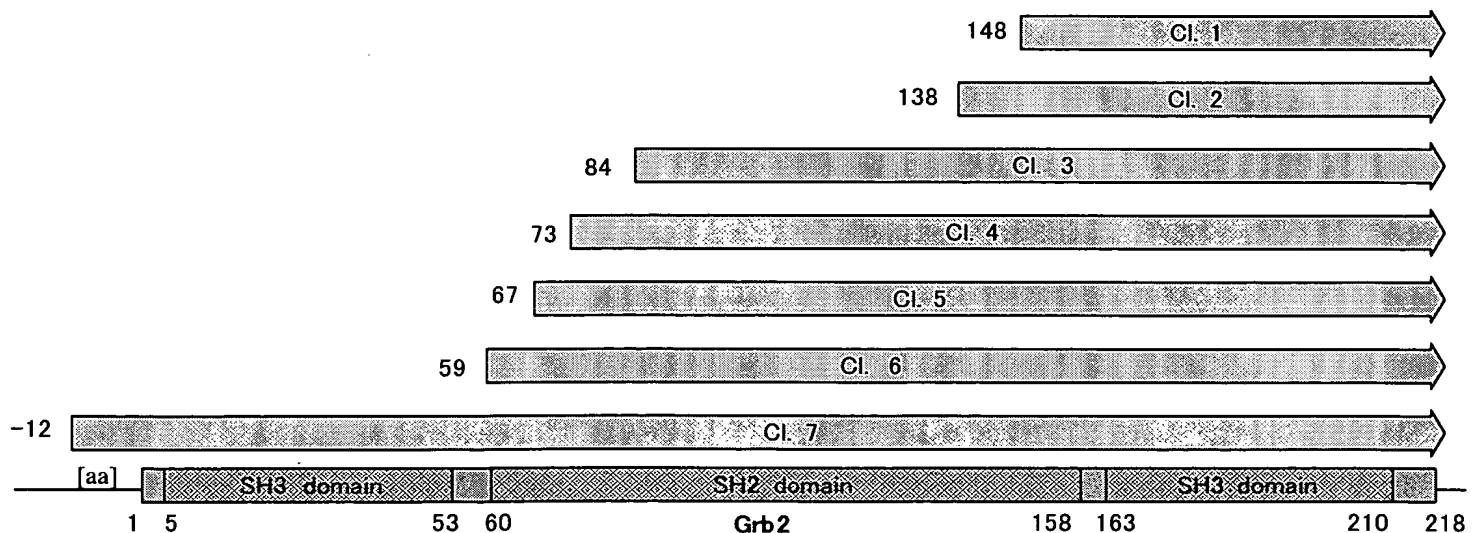


## Grb2のドメインスクリーニング

### • Grb2

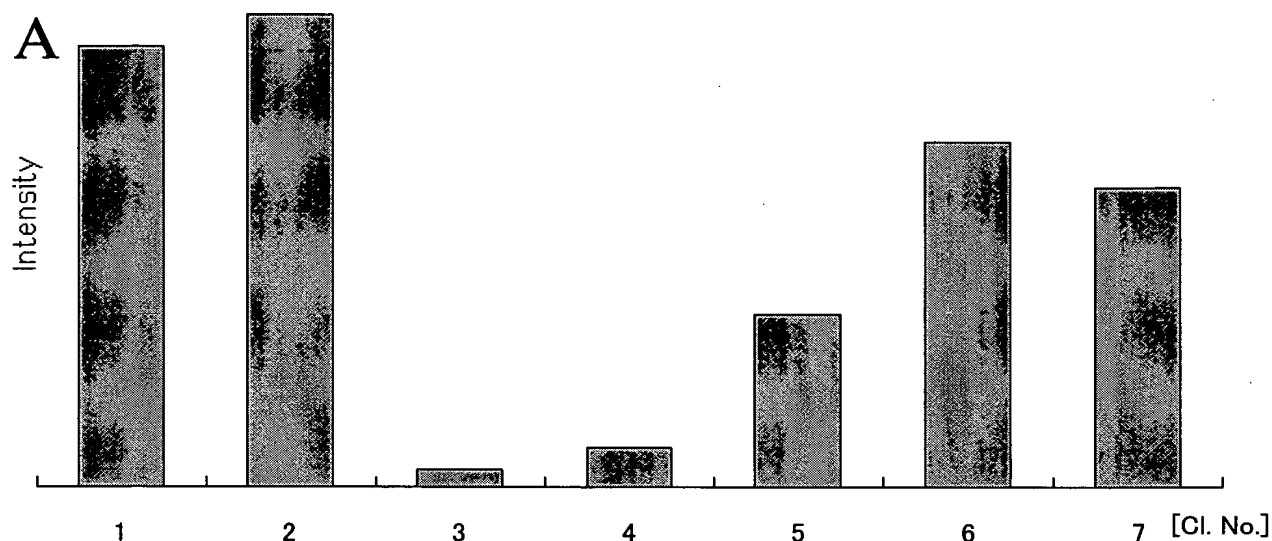
- ・全長で可溶性 GFP-fusionは発光
- ・SH3 SH2 SH3ドメインを含む
- ・立体構造がすでに解析されている  
(Maignan, S. et. Al. Science 268, 291-293 1995)

## スクリーニングにより選択された Deleted cloneの位置

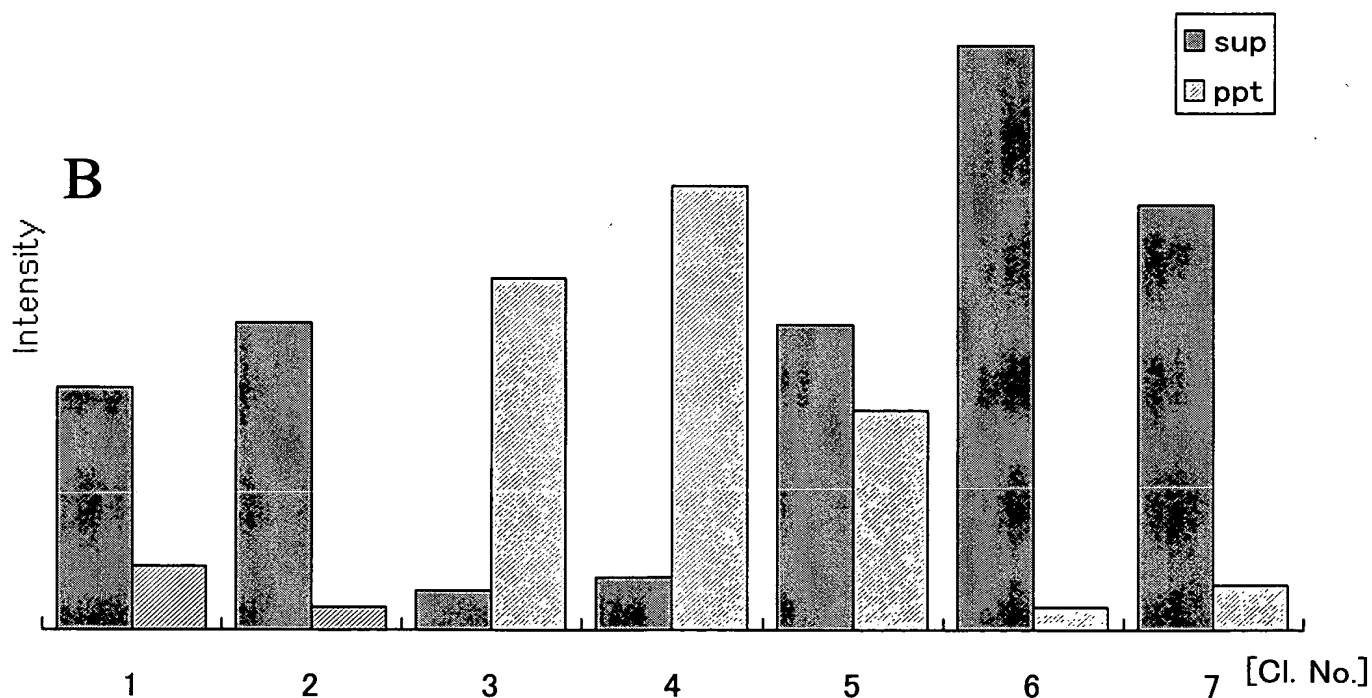


\*ここでのドメインはアミノ酸配列の保存性から導き出された位置を示している

# Grb2のDeleted cloneの解析 (1)



**A 無細胞合成による GFP-fusionの蛍光強度**  
(励起 485nm 発光 535nm)



**B 無細胞合成による単体での可溶性**  
( $^{14}\text{C}$ , scintillation counter)

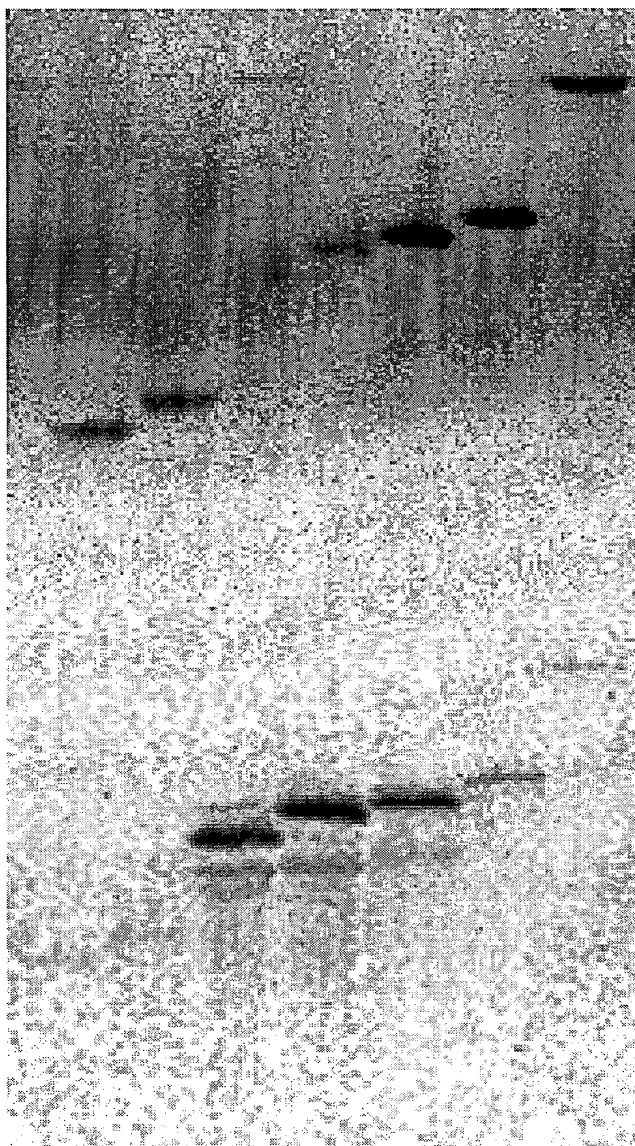
## Grb2のDeleted cloneの解析(2)

C

1 2 3 4 5 6 7

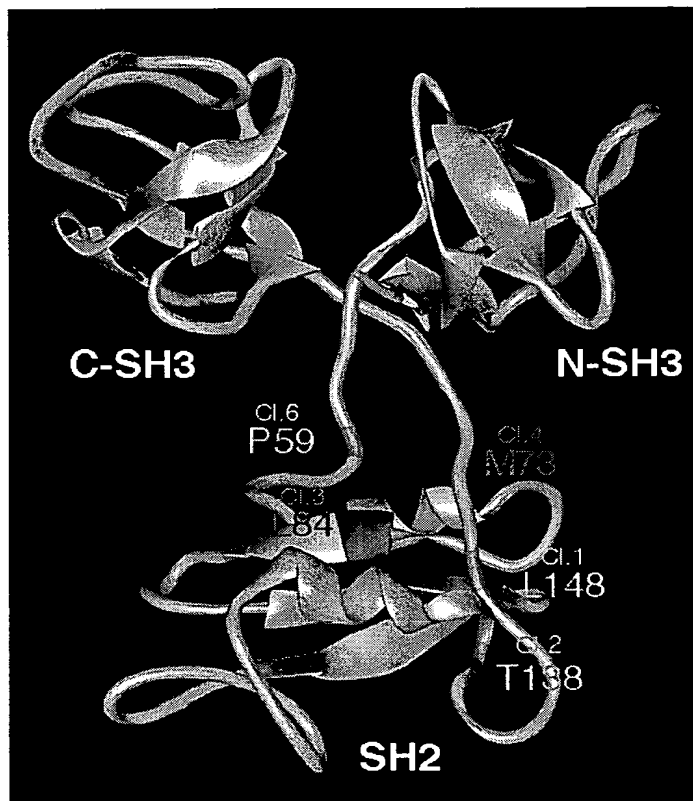
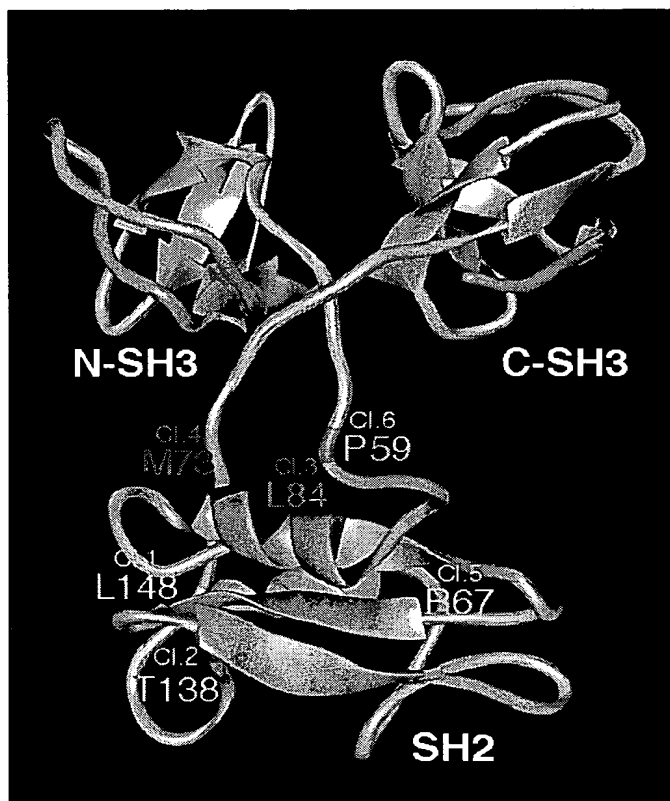
sup.

ppt.



C. 無細胞合成による単体のSDS-PAGE  
(<sup>14</sup>C,MacBas)

# 立体構造とDeletion部位の比較



■ : Soluble

■ : Partially soluble

■ : Insoluble

## 結論

- Exonuclease IIIを用いることで、均一でサンプリング範囲の広いDNA randomly deleted library を作成することが可能になった。
- DNA deleted libraryとタンパク質C末端側でのGFP-fusionを組み合わせ、さらに、大腸菌と無細胞系を用いて2段階でスクリーニングすることにより、可溶性の高いドメインを、効率よく選び出すことが可能になった。
- Grb2においては、本系が導き出した可溶性の高いfragmentの境界と、立体構造上のドメイン境界との間に強い相関があった。

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**